Applications of Physics-Informed Neural Networks in Power Systems - A Review

Bin Huang, *Student Member, IEEE*, and Jianhui Wang, *Fellow, IEEE*

***Abstract*—The advances of deep learning (DL) techniques bring new opportunities to numerous intractable tasks in power systems (PSs). Nevertheless, the extension of the application of DL in the domain of PSs has encountered challenges, e.g., high requirement for the quality and quantity of training data, production of physically infeasible/inconsistent solutions, and low generalizability and interpretability. There is a growing consensus that physics-informed neural networks (PINNs) can address these concerns by integrating physics-informed (PI) rules or laws into state-of-the-art DL methodology. This survey presents a systematic overview of the PINN in the domain of PSs. Specifically, several paradigms of PINN (e.g., PI loss function, PI initialization, PI design of architecture, and hybrid physics-DL models) are summarized. The applications of PINN in PSs in recent years, including state/parameter estimation, dynamic analysis, power flow calculation, optimal power flow, anomaly detection and location, and model and data synthesis, etc., are investigated in detail, followed by the summary and assessment of relevant works so far. Revolving around the characteristics of PSs and the state-of-the-art DL techniques, this paper outlines the potential research directions and attempts to shed light on the deeper and broader application of PINN on PSs.**

***Index Terms*—Deep learning, first principle, neural networks, physics-informed neural networks, smart grids.**

I. Introduction

THE advances of deep learning (DL) techniques in recent years have aroused great interest in academia and industry. The deeper architectures result in a better representation learning and data possessing capability. The phenomenal performances of deep neural networks (DNNs) facilitate the development of data-driven solutions to those physics-related problems in which the physical mechanism is not fully understood, or in which it is intractable to derive a high-resolution numerical solution with limited computational resources. The fast feed-forward feature, the black-box model, and the powerful function approximation ability are the primary reasons for the popularity of the DNNs in these fields.

Despite all these great advantages, DNN still cannot be regarded as a panacea to problems in the real world. The applicability of the DNNs has encountered certain challenges in physics-related or engineering fields [1]. First, the training of the DNNs often requires a great quantity of high-quality training data to guarantee performance. However, training data is very scarce or costly to obtain in many scenarios. Second, the uses of the DNN are also plagued by the black-box nature of DNNs and the lack of theoretical analysis of the lower bound of performance, resulting in doubts about the reliability of the results, especially in safety-critical applications. Third, the plain DNNs are likely to produce physically infeasible results without a delicate design and domain-specific knowledge. Finally, it is intractable to provide a theoretical guarantee on the generalizability of the generic DNNs [2], which means that the DNNs cannot guarantee consistent out-of-sample performance.

As one of the paths to resolve these issues, linking physics-based prior knowledge to the development of the DNNs has received increasing attention in the research community. This research direction is named after physics-guided neural networks, physics-aware neural networks and physics-informed neural networks (PINNs), etc. We will use the term PINN throughout the survey in order to be consistent. Different from the mainstream paradigm in the DL community that focuses on purely data-driven approaches, PINN is characterized by exploiting the scientific knowledge or the physics laws to guide the optimization, architecture design, and implementation of the DNNs [3]. The core paradigms of the PINN so far include (i) physics-informed loss function, (ii) physics-informed initialization, (iii) physics-informed design of architecture, and (iv) hybrid physics-DL models. The following potential advantages distinguish and motivate the development of PINN: (i) improved prediction accuracy of certain applications; (ii) improved interpretability; (iii) physically consistent results; (iv) better generalizability; (v) shrunk search space of the weights; improved training efficiency; better convergence performance; and (vi) enhanced sampling efficiency; reduced reliance on a large amount of training data.

Although the concept of PINN has not been put forward for a long time, due to the potential superiority of PINN, it has been preliminarily investigated and explored in many disciplines, including hydrology [4], chemistry [5], material [6], Earth systems [7], and hydromechanics [8], etc. As power systems (PSs) are large-scale artificial complex physics systems that are crucial to human society, PSs are well-known for their highly dynamic and non-linear operating characteristics.

Manuscript received 18 August 2021; revised 27 December 2021 and 1 March 2022; accepted 19 March 2022. Date of publication 25 March 2022; date of current version 22 December 2022. This work was supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office Award Number DE-EE0009337. Paper no. TPWRS-01329-2021. *(Corresponding author: Jianhui Wang.)*

The authors are with the Electrical and Computer Engineering Department, Southern Methodist University, Dallas, TX 75205 USA (e-mail: bin@smu.edu; jianhui@smu.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPWRS.2022.3162473.

Digital Object Identifier 10.1109/TPWRS.2022.3162473

Thus, numerous challenging tasks stem from developing the control schemes for such sophisticated systems. Due to the rapid penetration of advanced sensors, smart meters, parallel computing resources/frameworks, and the advanced representation learning ability of the DNNs, recent advanced DL techniques have intensive applications in the field of PSs [9]. As PSs are members of diverse physical systems, the concept of PINN can be naturally transplanted to PSs, which is one of the most promising future research paths of the DL on PSs.

Although there are some literature surveys and summaries of the PINNs, these reviews either focus on the general framework of the physics-informed machine learning (ML) techniques and their applications in the general knowledge field [3], or analyze the applications in some specific fields such as chemistry [10] and cyber-physical systems [11]. Different from the excellent prior works on [3], [10], [11], this review attempts to conduct a relatively exhaustive survey of PINNs on PSs with cutting-edge developments (e.g., state/parameter estimation, dynamic analysis, power flow calculation, optimal power flow, anomaly detection and location, model and data synthesis, and so forth), particularly over the past two years. More importantly, respecting the operational characteristics and physical laws of PSs, this paper proposes specific potential research directions for PINN on PSs.

The remainder of this survey is organized as follows. Section II introduces several paradigms of PINN. Section III exhibits a comprehensive review of PINN on PSs with the cutting-edge developments at the application level, followed by consolidating the state of the art and pointing out the potential research directions. A broad and general vision for the future development of PINNs on the open problems in PSs is provided in Section IV. Section V presents the concluding remarks.

II. Overview of PINN Paradigms

This section provides details for four types of paradigms of implementing the PINN. Specifically, for each paradigm, we will introduce the general idea of how to integrate physics information into the DNNs, discuss the corresponding benefit brought by the integration, and present some relevant examples and references. Fig. 1 outlines these four paradigms and highlights their positions on the implementation of DNNs.

A. Physics-Informed Loss Function

There are two key steps in implementing the physics-informed loss function: 1) assigning physical meanings to the variables in the output or hidden layers of neural networks (NNs); 2) augmenting the governing equation involving these physical variables to the loss function through the regularization. This paradigm is presented in part 1 in Fig. 1.

The loss function of the PINN $\mathcal{L}$ can be formulated as:

$$\mathcal{L} = L(\hat{\mathbf{y}}, \mathbf{y}) + \lambda R(\mathbf{W}, \mathbf{b}) + \gamma R_{\text{Phy}}(\mathbf{X}, \hat{y}) \tag{1}$$

where L is the conventional loss function which measures the distance between the predicted output and the target output in a supervised learning manner; R is the parametric regularization term (e.g., L1 and L2 norm regularization) imposed on the DNN

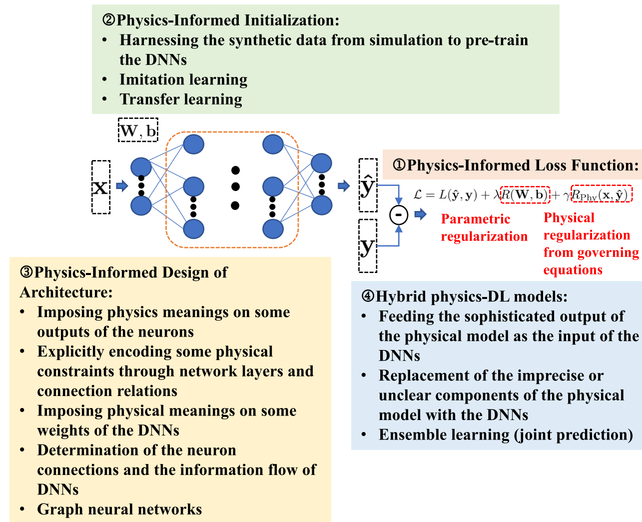


Fig. 1. Overview of physics-informed neural network paradigms.

weights $\mathbf{W}$ and bias $\mathbf{b}$; R_{Phy} is the physical regularization term (PRT) based on $\hat{y}$ and the equations of physical principles. λ and γ are regularization hyperparameters.

The PRT distinguishes the PINN from the vanilla NN. The PRT can take a variety of forms, and the design of the PRT can be based on the engineering problem to be solved. To make PRT more intuitive, we give the guiding example of solving partial differential equations (PDEs) using the PINN for capturing rotor angle δ and frequency $\omega := \frac{\partial \delta}{\partial t}$ [12], [13] in the single machine infinite bus system (SMIBS), of which the swing equation is a PDE:

$$\xi \frac{\partial^2 \delta}{\partial t^2} + \kappa\omega + BV_g V_e \sin(\delta) - P = 0 \tag{2}$$

where ξ is the inertia constant, κ denotes the damping coefficient, B represents the susceptance between the generator and external grid, V_g and V_e are the voltage magnitudes of generator and external grid, respectively, and P is the mechanical power of the generator.

In a general case, since a NN is a universal function approximator, it can be used to approximate an unknown function $u(t, x)$ associated with time variable t and input variable x. In the SMIBS case, u can be used to refer to δ associated with t and P:

$$u(t, x) := \delta(t, P) \tag{3}$$

In a general case, without loss of generality, u is subject to a PDE [12]:

$$\frac{\partial u(t, x)}{\partial t} + \mathcal{N}[u] = 0, x \in \Omega, t \in [0, T] \tag{4}$$

where $\mathcal{N}$ denotes a nonlinear differential operator which describes the dynamics of the physical system; Ω and T constrain the range of x and t, respectively, and are based on the actual operating conditions of the physical system. In the SMIBS case, (4) corresponds to the swing equation in (2).

In a general case, the PDE in (4) can be approximated as a PINN $f(t, x)$:

$$f(t, x) = \frac{\partial u(t, x)}{\partial t} + \mathcal{N}[u] \tag{5}$$

where $\frac{\partial u(t,x)}{\partial t} + \mathcal{N}[u]$ can be obtained from taking the derivative of the output of the NN $u(t, x)$ w.r.t t/x and adding some constant terms. Note that the NNs f and u can share exactly the same parameters and architecture based on the auto-differentiation features of the popular DL packages. In the SMIBS case, the PINN is built on (2) and (5):

$$f(t, P) := \xi \frac{\partial^2 \delta}{\partial t^2} + \kappa\omega + BV_g V_e \sin(\delta) - P \tag{6}$$

In a general case, denote the initial and boundary training data of $u(t, x)$ as $\{t_u^i, x_u^i/\overline{u}^i\}_{i=1}^{N_u}$, where i and N_u are the index and number for the training data samples, respectively; t_u^i and x_u^i are the input data; $\overline{u}^i$ is the output data. In fact, to approximate $u(t, x)$, we can feed $\{t_u^i, x_u^i/\overline{u}^i\}_{i=1}^{N_u}$ to NN and define the mean square error (MSE) loss function as follows to train it:

$$\mathcal{L}_{MSE,u} = \frac{1}{N_u} \sum_{i=1}^{N_u} \left| u\left(t_u^i, x_u^i\right) - \overline{u}^i \right|^2 \tag{7}$$

MSE is a common loss function for regression problems in DL, and it is defined as the mean of the squares of the error between the predicted value and the ground truth [14].

However, the effects of direct usage of (7) are not satisfactory. This is because in the real world, the mechanism behind $u(t, x)$ is often very complex, and in order to approximate $u(t, x)$ as closely as possible, the above method requires the training data to be nearly perfectly distributed in the input space. A large amount of training data is also needed to guarantee sufficient generalizability. These hurdles motivate the introduction of PINN.

Denote a finite set of collocation points of $f(t, x)$ as $\{t_f^j, x_f^j\}_{j=1}^{N_f}$, where j and N_f are the index and number for the collocation points, respectively. Similar to (7), the MSE loss function can be derived to train the PINN $f(t, x)$:

$$\mathcal{L}_{MSE,f} = \frac{1}{N_f} \sum_{j=1}^{N_f} \left| f\left(t_f^j, x_f^j\right) \right|^2 \tag{8}$$

Considering that $f(t, x)$ and $u(t, x)$ share a set of NN parameters and architecture, they can be jointly trained via the following loss function:

$$\begin{aligned}\mathcal{L}_{MSE} &= \mathcal{L}_{MSE,u} + \mathcal{L}_{MSE,f} \\ &= \frac{1}{N_u} \sum_{i=1}^{N_u} \left| u\left(t_u^i, x_u^i\right) - \overline{u}^i \right|^2 + \frac{1}{N_f} \sum_{j=1}^{N_f} \left| f\left(t_f^j, x_f^j\right) \right|^2\end{aligned} \tag{9}$$

where the second term $\mathcal{L}_{MSE,f}$ in $\mathcal{L}_{MSE}$ is the PRT. $\mathcal{L}_{MSE,f}$ represents the degree of violation of the governing PDE in (4). In the SMIBS case, training data is defined as:

$$\{t_u^i, x_u^i/\overline{u}^i\}_{i=1}^{N_u} := \left\{t_u^i, P_u^i/\overline{\delta}^i\right\}_{i=1}^{N_u} \tag{10}$$

where t_u^i is sampled from the time steps in the system evolving interval, and P_u^i is sampled from the range of generator output. Likewise, the definition of collocation points in the SMIBS case is similar to training data except that the need for sampled δ is eliminated. Interested readers can refer to [12], [13] for further details.

The fusion of the PRT brings the following benefits. First, the PRT is necessary to derive a feasible numerical solution in some applications, such as solving nonlinear PDEs. Second, the incorporation of $\mathcal{L}_{MSE,f}$ encourages the NN to generate the physically consistent solution u. Interpretability is a scarce feature in NN, whereas it is helpful for debugging and for users of NN to interpret the output of NN. Third, the search space of the parameters $\mathbf{W}$ and $\mathbf{b}$ is shrunk significantly, and thus the demand for the number of training samples can be reduced. In addition to the above PRT for solving PDEs, other exemplary PRTs designed for PSs will be introduced in Section III.

B. Physics-Informed Initialization

The optimization of DNNs is highly nonlinear and nonconvex. Such an optimization problem is sensitive to the selection of the initial solution. Hence, the initial weights of DNNs have a great influence on convergence performance. Good initial weights can accelerate the convergence and prevent the optimization from getting stuck in the local minimum. The parameter initialization of DNNs in the engineering domain can be inspired by the first principles and is denoted by physics-informed initialization in this paper. The physics-informed initialization can achieve similar results to transfer learning. This paradigm is presented in part 2 in Fig. 1.

One approach to implement the physics-informed initialization is harnessing the synthetic data from the simulation to pre-train the DNNs and then fine-tuning them with fewer observation data from the real world. [4] employs this strategy to simulate the lake temperature dynamics. The PINN is trained on the simulated dataset generated by the physical model first. Then the weights are transferred to the lake temperature prediction task and are trained using the real observation data. [15] reports that the self-driving algorithm can obtain a decent initial policy by pre-training in a simulator built on a video game physics engine. Another crucial area of application of the physics-informed initialization is imitation learning, which is detailed in Section III-D along with the relevant applications in PSs.

Overall, physics-informed initialization has similar advantages to transfer learning, which can avoid learning from scratch and improve training efficiency. This scheme is particularly useful for situations where the training data is scarce, or data acquisition costs are high.

C. Physics-Informed Design of Architecture

Since traditional DNNs are model-agnostic, there are numerous redundant connections and parameters that hinder their out-of-sample performances. The idea of physics-informed design of architecture is to reshape the architectures of DNN with inspiration from the physical and engineering domain. This paradigm is presented in part 3 in Fig. 1.

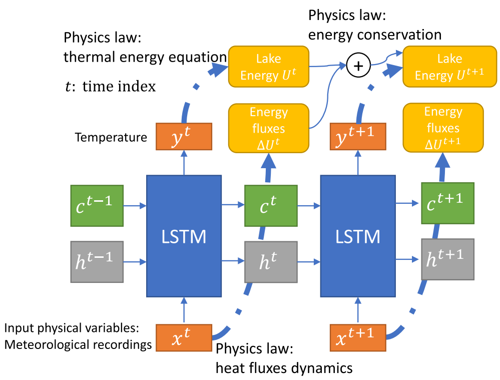


Fig. 2. The information and energy flow of the physics-informed RNN model for predicting temporal lake temperature. Three physics laws, i.e., thermal energy equation, heat fluxes dynamics, and energy conservation, are encoded explicitly.

One way to leverage the physics information to guide the architecture is imposing some physical meanings on some outputs of the neurons in the hidden layers. Then, specialized network layers and connection relations can be imposed on these neurons, thereby explicitly encoding some physical constraints. For example, [16] applies this strategy to the lake water modeling. A density monotonicity-preserving recurrent neural network (RNN) architecture is proposed for the spatial lake temperature prediction task. The density difference between different depths is derived by feeding hidden states of long short-term memory (LSTM) cells to the fully connected layers and the Rectifier Linear Unit (ReLU) activation function. The intermediate physical variable (i.e., density) is updated by combining the density difference and the density from the last state. The embedding of physical constraints alleviates the black-box nature of DNNs and makes parts of DNNs interpretable by users.

Apart from manipulating the neurons of NN, another viable approach is imposing physics meaning on some weights of the DNNs and keeping them fixed throughout the whole training process. An instance can be found in [17], where the weights of DNNs are set as the parameters of the seismic wave propagation governing equations. In this way, the physics constraints are integrated seamlessly, which leads to more robust training and better interpretability.

Domain knowledge can also be used to determine the node connections and guide the information flow of DNNs. This idea is well exemplified in Fig. 2, where the temporal evolution of lake temperature and energy flow is embedded in the RNN [4]. In addition to the regular RNN information flow, the output of the RNN is further tightly linked via encoding the law of energy conservation.

Another typical example of this paradigm is graph neural networks (GNNs). Thanks to the expressive power of GNNs and the intrinsic regularization from the mechanism of aggregating information from neighbors only, GNNs show superior performance across multiple areas and applications. GNNs are capable of processing non-Euclidean structured data, i.e., graphs with nodes and edges. Since the topology of PSs can be represented by graphs, GNNs applied to PSs can also be regarded as a type of topology-aware NNs. Since a comprehensive summary of GNNs in the field of PSs has been given in [18], this survey will only report the selective progress of GNNs on PSs in which physical principles are integrated in a more sophisticated manner other than the pure data-driven way. Interested readers can refer to [18] for further details.

D. Hybrid physics-DL Models

Hybrid physics-DL models can be distinguished from the other methods mentioned above by the explicit coexistence of the physics and DL models. From the perspective of mathematical modeling, according to the completeness of a priori information, the physical model can be referred to as the white-box model, and the DL model can be referred to as the black-box model. Due to the complexity of engineering problems in the real world, it is often unrealistic or cumbersome to derive the perfect first principle equation and specific parameters. As a compromised version of black-box and white-box models, the hybrid physics-DL model (also known as a grey-box model [19], [20]) can make full use of a priori knowledge and the function approximation capability of DL simultaneously. Hence, it is both first-principle-based and data-driven. This paradigm is presented in part 4 in Fig. 1.

In general, there are three ways to implement hybrid physics-DL models. The first method is physics-informed feature engineering, which refers to leveraging the insights of prior physical knowledge to select, transform, and synthesize the most informative input variables from raw data for the ML model. A demonstrative example in PSs is [21], where ensemble decision trees are harnessed to predict online rotor angle stability with high wind power penetration. Other than feeding the unprocessed data (e.g., rotor angles, speeds, terminal voltages of synchronous generators) directly obtained from measurement devices to an ML model, a novel rotor angle stability index (SI) is extracted from the kinetic and potential energies of a system after fault clearance and serves as the input feature. The ML model built upon SI is more accurate than the existing ML-based stability prediction methods because SI is more discriminative and recognizable, especially under the high wind penetration level. The second method is the hybrid learning of black-box and white-box models. In hybrid learning, NN and physical models can be each other's input and output, and the physical model will affect the loss function and gradient of NN. Alternatively, the imprecise or unclear components of the physical model can be replaced by the DL model. One example can be found in [22], where a DNN is employed to estimate unknown variables in the turbulence closure model to compensate for omissive physical details. The third method is inspired by the idea of ensemble learning, that is, concatenating the outputs of the physical model and the DL model to make a joint prediction [23]. The rationale behind this method is that the classical physical model is better at modeling the physical process with a long-term scale, whereas the DNNs are better at capturing the hidden representation behind the short-term stochastic interactions. A demonstrative example can be found in [24], where a hybrid data-driven physics model-based false data injections detector is presented. The result of the physical bad data detection based on the Chi-Square test of state estimation (SE) is fused with the diagnosis of the ML detector called Correlation-based Detection through normal normalization to render a complementary attack-detection solution.

III. Applications in Power Systems

In recent two years, the application of PINN in PSs has gradually surfaced, which is summarized in Table I. Table I organizes the areas, specific applications, integration modes, baseline methods, and qualitative and quantitative effects of PINNs. These applications involve state/parameter estimation, dynamic analysis, power flow calculation, optimal power flow, anomaly detection and location, model and data synthesis, and so on. Here are some interpretations of the baselines/effects column in Table I. *1) Quantitative comparison:* All baselines are compared with PINNs. The data presented in the table shows the amplitude of performance of PINNs higher (↑) or lower (↓) than baselines. The symbol "x" means times, and the symbol "~" denotes about/approximately. For references that do not highlight the quantitative comparison, the qualitative effect is provided instead. *2) Selection of baselines:* Baselines in PINN's papers can be classified into two major categories: pure data-driven approaches and pure model-based approaches, respectively. For each major category containing more than one baseline method, not all baselines will be reported due to the limitation of space. Instead, the performance of the best baseline method among multiple baselines in this major category will be shown. For instance, in [48], authors compare the PINN method to linear regression (LR), support vector regression (SVR), k-nearest neighbors, and NN methods, which all belong to pure data-driven approaches. SVR and NN are always superior to the remaining two methods under all cases among these four baselines. Hence, the performance of SVR and NN are reported. Since none of the pure model-based approaches are compared, the performance of this major category is unfilled. *3) Selection of cases:* Sometimes, there may be multiple case studies or test systems in one reference. Then the performance of PINNs in multiple cases may appear in the form of intervals (e.g., 3%-14%). In some literature, the results on more realistic cases (considering measurement noise, limited observation, etc.) or test systems (real power grid) are reported preferentially. *4) Efficiency:* The efficiency exhibited in the table refers to the calculation time in the deployment or testing phase if not otherwise specified.

This section will review the details and motivations behind the relevant literature, and point out potential research directions, revolving around the characteristics and pitfalls of applications.

A. State/Parameter Estimation

1) Literature Review: Power system SE is a vital function in the Energy Management System. The basic principle of SE is to estimate the state of PSs using the redundancy of measurement systems and power flow equations (PFEs) under noisy measurements.

Given the inherent nonconvexity, previous SE schemes, e.g., optimization methods and weighted least square (WLS), are either sensitive to solution initialization, or time-consuming due to iterative calculation. DNN is a promising tool to bypass these hurdles because the DNN, which is trained to capture the patterns in the SE offline, can rapidly output the estimation results based on efficient feedforward matrix calculation in the testing phase. A DNN with a skip-connection architecture, named prox-linear net, is proposed in [25] to mimic a SE algorithm called prox-linear solver. The skip-connection architecture is inspired by the loop nature of the prox-linear solver and enables efficient training. The proposed novel prox-linear net outperforms the widely adopted Gauss-Newton SE solver. Building on [25], [26] attempts to imitate the alternating minimization solver by constructing a Gauss-Newton unrolled NN with skipping-connections. The topology-aware priors are incorporated into DNN through a GNN. Impulsive communication noise and cyberattacks can corrupt the SE results in the form of bad data, rendering grossly biased estimation results. To cope with these stochastic adversaries, a distributionally robust training scheme is employed to robustify the PINN, yielding stable estimation relative to competing alternatives.

Ref. [27] focuses on the limited observability of the distribution grid, where the measured quantity is less than the number of variables to be estimated. The estimation process is formulated as a voltage prediction problem in [27]. The DNN is employed to approximate the mapping between the voltage phasors of the previous time steps under the fully observable scenario and the voltage phasors at the current moment, where only partial system measurements are observable due to equipment failure or malicious attacks. The complex linear PFE based on the nodal admittance matrix is explicitly encoded into the loss function to guide the training of DNN. Case studies show that reducing the degree of observability by 40% tends to decrease the performance of the PINN by about 10%. Closely related to [27], [28] exploits the approximate separability physical property of the SE of the distribution grid. The separability property inspires the design of a topology-aware DNN via optimal placement of phasor measurements units (PMUs). Specifically, neurons in NNs are considered as nodes in PS, and the connection between neurons follows the principle of localization, which means neurons representing two different partitions cannot have a direct connection. This strategy can be interpreted as the hard-coding of weights matrix. With the advantage of alleviating over-fitting, the proposed approach shows robust performance even under noisy/corrupted measurements and pseudo-measurements. However, given that DNNs have a multi-layer structure, the scheme of assuming the neurons in each layer of NNs as nodes in PSs is prone to have scalability issues when it comes to large-scale systems. Furthermore, the connection between consecutive layers in [28] still exhibits redundancy. [29] takes a step further, and zeros out the redundant layer connections considering the diameter of the partitions. The diameter is defined as the longest distance between any two nodes in the graph. This strategy enhances the generalizability of the PINN and reduces the training burden. The test results indicate that the decrease or even loss of pseudo measurements has little impact on the performance of PINN.

Ref. [31] and [32] approach the joint estimation of state and parameter problem via the Power-GNN, where the loss function

TABLE I
APPLICATION OF PHYSICS-INFORMED NEURAL NETWORK ON POWER SYSTEMS

Area	Specific application	Integration mode	Baselines / Effects
State/ parameter estimation	Real-time SE [25]	Architecture design (skip-connection)	Gauss-Newton / accuracy ↑ -9%-37% / efficiency ↑ 140x-550x NN / accuracy ↑ 45%-82% / efficiency ↓ 3%-14%
	Real-time transmission SE [26]	Architecture design (skip-connection)	[25], NN / accuracy ↑ / against adversarial attacks
	Distribution SE [27]	Loss function (incorporate PFE)	WLS / accuracy ↑ ~10^6x under partial observability NN / accuracy ↑ ~10x under partial observability
	Distribution SE [28], [29]	Architecture design (prune neuron connection based on topology)	WLS / accuracy ↑ 71% / efficiency ↑ 165x under corrupted measurements
	Parameter estimation [30]	Architecture design (incorporate PFE) Initialization	Maximum likelihood estimation / accuracy ↑ 14%-30% GNN / accuracy ↑ 5%-10% under noisy measurements
	Joint parameter and state estimation [31], [32]	Loss function (incorporate PFE)	GNN / accuracy ↑ 29%-2 × 10^3x under limited observability
	Transmission SE [33]	Hybrid learning Architecture design (PFE-based decoder)	WLS / voltage accuracy ↑ 54%-66% / angle accuracy ↑ 0%-37% WLAV / accuracy ↑ $5\% - 20\%$ with corrupted data
	DSE [34]	Architecture design (automatic differentiation) Loss function	UKF / accuracy ↑ ~10x with standard and fast systems; accuracy ↓ ~80x with slow systems
	DSE [35]	Hybrid models (hybrid learning)	extended KF / efficiency ↑ ~80x / accuracy ↑ 40%-85% under base fault / nonconvergence ↓ ~80%
Dynamic analysis	Identification of state parameters [13], [36]–[39]	Architecture design (automatic differentiation) Loss function (incorporate swing equation)	$ode45$ solver / efficiency ↑ 28x-87x / small samples needed (40) [13] RK / efficiency ↑ 100x-1000x / No training data needed [36], [37] NN / accuracy ↑ 10%-75% [39]
	Converter dynamics [40]	Architecture design Loss function	No time-domain simulations needed
	Frequency regulation [41]–[43]	Hybrid models (Koopman [41], [42], Strategic utility function [43])	EDMD / accuracy ↑ ~10x [41] Actor-critic / smaller frequency fluctuation [43]
	Rotor angle stability [44], [45]	Hybrid models (Koopman [44], hybrid controller [45])	EDMD / better angle and speed prediction [44] PSSs / damping ratios ↑ 31% [45]
	AGC/LFC [46], [47]	Initialization [46] Hybrid models (emotion [47])	Evolutionary algorithms / regulation cost ↓ 0.4% [46] Proportional–integral controller / frequency deviation ↓ 37-59% [47]
Power flow calculation	Voltage prediction [48]	Architecture design (imitate PFEs)	SVR / accuracy ↑ ~2x / generalizability ↑ ~3x NN / accuracy ↑ ~4x / generalizability ↑ ~2x
	PF [49]	Architecture design (GNN)	NN / test loss ↓ 0.25x-120x / number of parameters ↓ ~1.6x-~26.3x / convergence speed ↑ 83x
	PF [50], [51]	Architecture design (GNN)	DCPF / accuracy ↑ ~5x-~100x / efficiency ↑ ~10x Newton-Raphson / efficiency ↑ ~140x-~200x
Optimal power flow	ACOPF [52]	Loss function (Lagrangian relaxation)	DCOPF / accuracy ↑ 10^2x-10^3x / efficiency ↑ ~10^3x
	DCOPF [53]	Architecture design Loss function (KKT conditions)	NN / optimality ↑ 5x / feasibility ↑ 4.3x / infeasibility ↓ 15.6x
	Smart inverter control [54]	Architecture design (automatic differentiation) Loss function (Sensitivity consistency)	NN / testing error ↓ ~10x-10^2x / accuracy ↑ ~10x-10^3x / training data amount ↓ ~4x
	OPF [55], [56]	Loss function (KKT conditions)	NN / training time ↑ 3x / test error ↓ 20%-30% / worst-case violations ↓ 11%-63%
	Inverter control [57]	Architecture design (communication design)	Stochastic programming / near-optimal behavior
	Volt/Var control [58], [59]	Hybrid model (action-value function)	Optimization / No voltage deviation / power loss ↓ 1.0%-2.7%
	Volt/Var control [60], [61]	Hybrid model (surrogate model)	Load flow-based environment / near-optimal behavior / No power flow calculation needed
	Restoration [62] Control [63]	Initialization (imitation learning)	DRL / unsolvable cases ↓ 4%-13% / three epochs to find optimal policy [63]
Anomaly detection and location	High impedance faults detection [64]	Loss function (elliptical regularization)	AE / accuracy ↑ ~15% PCA / accuracy ↑ ~30% ER / accuracy ↑ ~39%
	Fault location [65], [66]	Architecture design (GNN)	NN / accuracy ↑ ~30% under 15% label rate GCN / accuracy ↑ ~35% under 15% label rate
Synthesis	Model synthesis [67], [68]	Hybrid models	random-matrix / success rate ↑ 27% [67] GMM / accuracy ↑ ~2x-~7x [68]
Miscellany	Wind power estimation [69]	Architecture design (imitate wake model)	GNN / error ↓ 24% / better interpretability

is built upon the residuals of PFEs. The limited observability issue is addressed in [31] by eliminating unobservable nodes of the PFEs using Kron reduction. [32] differs from [31] in that [32] explicitly incorporates the Newton Raphson scheme into the loss function. The results show that Power-GNN enables the reconstruction of parameters and state estimation simultaneously under limited observability, where physics-blind methods may fail. [31] also reports the demonstration on the continental European grid. Different from most studies focusing on single-phase circuits, [30] proposes the graphical learning method to estimate the line parameters of a three-phase unbalanced distribution network, i.e., impedance matrix, using the data from smart meters. The line parameters are embedded into a GNN as the trainable parameters. Specific components of the GNN are replaced by hard-coded three-phase PFEs-based transition functions. Besides, available prior distributions of the line parameters from Global Positioning System are used to initialize and constrain the parameters to be identified. The superiority of the proposed method becomes more pronounced as the noise level of smart meter measurement increases.

A hybrid learning autoencoder framework is proposed in [33] to capture the temporal correlations of system states and conduct the SE of the transmission network. The proposed autoencoder features the alternating current PFE-based (ACPFE) decoder, where the intermediate variables (measurements magnitude and angle) from the encoder are mapped to estimated measurement data. The weights of the encoder are updated based on the back-propagation of the cumulative error between actual and estimated measurement vectors. Case studies in [33] show the superior performance of PINN over WLS and the least absolute value (LAV) estimator. Under the corrupted data scenarios (e.g., contaminated data, missing data, and miscommunication), PINN keeps outputting solid results, while traditional methods have encountered the non-convergence issue.

In the domain of dynamic state estimation (DSE), Kalman filter (KF) and its derivatives, e.g., unscented Kalman filter (UKF), are the most common algorithms. However, the significant nonlinearity and the challenging initialization cannot be handled well in the above methods. To overcome these challenges, [34] investigates the performance of PINNs on characterizing the frequency dynamics of PSs. Leveraging the automatic differentiation of the NN framework, the swing equation describing the dynamic characteristics of the system is represented by a DNN, where the output neurons are estimated as the rotor angles and generator speeds. The loss function consists of the physical regularization term from the swing equation and the quadratic loss term from the real measurement data. Two interesting results reported in the case studies are 1. the magnitude and type of noise have a minor impact on PINN, which demonstrates the usefulness of physics regularization; 2. PINNs encounter vanishing gradient issues on systems with slow dynamics. [35] leverages the hybrid learning of black-box and white-box models to approach the online DSE of the rotor angle and speed of synchronous generators. The hybrid-learning framework can be viewed as an autoencoder (AE) with the decoder replaced by the physical model. Dual NNs, i.e., rotor angle and speed estimator NNs, produce estimated rotor angle and speed. The estimated rotor angle is fed into the physical measurement model that outputs estimated voltage measurements, followed by constructing reconstruction loss. On the other hand, the estimated speed is fed into the physical transition model, along with the prior estimated rotor angle, to generate a posterior estimated rotor angle. The difference between the prior and posterior estimated rotor angle forms the loss function for the speed estimator NN. The PINN based on the hybrid learning shows superior computational efficiency, comparable accuracy, and best convergence performance under faulted and noisy measurement relative to the variants of the KF.

2) Summary and Potential Research Directions: Due to the data-intensive and error-tolerant nature of SE for power grids, PINN currently has a wide range of applications in the field of state and parameter estimation. General physics knowledge is fused in diverse ways. As a bridge between purely data-driven approaches and physical model-based approaches, PINN shows significant competitiveness in terms of estimation accuracy and robustness.

There are certain potential research directions to explore in this research area. 1) For example, the SE of large-scale three-phase unbalanced distribution networks is intractable and complex due to nonlinearity, nonconvexity, and coupling between phases. Currently, proposed methods face the dilemma between model complexity and computational efficiency/solvability. PINN will be a powerful tool for solving this problem with its superior generalization and nonlinear approximation capabilities. How to reflect the unbalanced components and the coupling relations between phases of unbalanced distribution networks in the design of DNN remains an exciting challenge. 2) The SE of the integrated energy systems (IESs), e.g., combined heat and power systems and integrated gas and electricity systems, is another promising application field of PINN. From the perspective of static SE, the line-pack and inertia of the heat/gas pipes are essential physical properties that can be integrated into the PINN [70]. The time-delay characteristics of the hot water or natural gas in the pipeline can be described by referring to the energy flow framework in [4]. From the perspective of dynamic SE of IESs, the PDEs that characterize the dynamics of the heat/gas network can be embedded in the PINN, in which the output neurons are regarded as the partial differentiation of state variables with respect to time or space. 3) The key to the hybrid learning method of black box and white box models is to devise the interaction mechanism between models, which is case-dependent. Considering that the parameters of some PS models (e.g., load models) are roughly calibrated due to the limited observations, an area worthy of further exploration is to generate additional data according to the output of the hybrid model to fine-tune the parameters of the physical model, rather than updating the parameters of DL as before.

B. Dynamic Analysis

1) Literature Review: Generally, PSs are non-linear dynamic systems, in which the dynamic transition of the state is described by non-linear PDEs. In order to ensure the stability of rotor angle and voltage, capturing PSs dynamics is one of the necessary measures. However, analyzing the dynamic evolution of PSs is notoriously computationally intensive. Traditional methods usually analyze the dynamics of PSs through computationally expensive time-domain simulations. As an alternative, Lyapunov function-based methods require less computational cost, whereas the conclusions drawn from these methods are conservative. The traditional DL algorithms are still limited by the lack of interpretability and high requirements for the training data [71]. The emergence of PINN provides an opportunity to solve the above challenges.

Inspired by [12] that proposes a generic PINN framework for solving PDEs, [13] tries to determine dynamic states, e.g., rotor angles and frequency, and to identify uncertain parameters such as inertia and damping using PINN. [13] adopts the same PINN structure as [12], where the PINN and the NN of the state variable share the same weights. The detail of the implementation is discussed in Section II-A, where the loss function is defined to incorporate the penalty upon the violation of the governing equation. The PINN in [13] features a small amount of labeled training data (only 40), resulting in 223 seconds of training time. In order to quantify the critical system dynamics indices in converters, the work exhibited in [40] leverages the PINN. The physical regularization term in the loss function is employed to assess the degree of satisfaction of the approximate states on governing differential equations. With the aim of capturing the dynamic behavior of power systems, [40] formulates an optimization problem where a group of differential and discrete constraints are incorporated to describe the system's state evolution. Subsequently, a PINN is formed by applying the automatic differentiation to the differential states. The loss function of this PINN is comprised of the reconstructed errors of both differential/algebraic states and governing equation violation terms. By carefully selecting the nonlinear activation function and applying the mixed-integer reformulation trick, the training procedure of the above PINN is converted to a mixed-integer linear programming (MILP), followed by solving by the off-the-shelf solvers. The presented results show the proposed PINN can produce an accurate approximation to dynamic response while eliminating exhaustive time-domain simulations. Another important observation is that the largest errors occur at the jumps in the algebraic variables. Ref. [36] extends the prior work in [13] to incorporate an implicit Runge-Kutta (RK) integration scheme, which can approximate the temporal evolution of a dynamical system. Another favorable feature of [36] is that it eliminates the use of the simulation data.

Ref. [37] is another extension to the initial approach in [13]. The data utilization efficiency is enhanced by introducing an additional residual of governing equations. The utilization efficiency of the dataset is elevated in [37] by including an additional loss term (i.e., dtNN). [37] develops a heuristic to balance the impact of multiple terms in the loss function of DNN, which are believed to have a huge influence on the accuracy. Inheriting the idea of [13], a similar PINN method is used in [39] to capture the temporal-spatial dynamical characteristics of battery energy storage systems (ESSs) using RNN. It is noted that the advantage of the PINN w.r.t the NN drops from 75% to about 10% as the prediction horizon increases.

How to analyze and control a nonlinear dynamic system has been an essential and challenging problem. In recent years, the development of the Koopman operator has provided a data-based approach to solve this thorny challenge: under appropriate assumptions, through the spectral decomposition of Koopman operators, we can find a set of functions as bases and perform coordinate transformation based on them. Under the updated coordinate system, the original system of interest becomes linear. In other words, the Koopman operator maps a finite-dimensional, potentially nonlinear dynamical system into an infinite-dimensional linear system. Koopman operator, as a powerful analysis tool, is being harnessed extensively in the domain of dynamic analysis of PSs: DSE [72], dynamics-aware economic dispatch [73], stability assessment [74]–[76], and coherency estimation [77], etc.

Despite some limitations, the most commonly used algorithm to implement the Koopman analysis is dynamic mode decomposition (DMD). Abundant nonlinear measurements need to be prepared for accurate analytics and predictions because the prior information of invariant subspace of state variables is not available [78]. In addition, DMD requires that the spectrum of the Koopman operator must be discrete [79], which also dramatically restricts the generalization ability and application of this method. The recent advances of PINN facilitate the derivation of the Koopman operator. In contrast to purely data-driven approaches, the constraints of physical principles are incorporated into network structure and cost function. The NNs only serve as effective nonlinear function approximators, and each step of the model has a corresponding physical explanation [79], [80]. For example, [79] devises a PINN schema, featuring the assignment of meanings to the inputs and outputs, e.g., time-series data and state variables at consecutive time steps. In addition, the loss function in [79] is comprised of 1) reconstruction error of AE; 2) the penalty item from the linearity constraints of the system after coordinate system conversion considering the system evolves linearly in the updated coordinate system; 3) the penalty item from the prediction of future state variables, which reinforces the linearity in the new coordinate system and drives the eigenfunctions and their inverses learned from the AE more stable. In the domain of PSs, [41] tries to learn the Koopman operator via an AE for model predictive control (MPC) of ESSs integrated with synchronous generators. A two-step training procedure is proposed to obtain the system matrix A, the control matrix B, and the network layer parameter Q separately. A and B are updated with Q fixed by solving a least square problem. In contrast, Q, which parameterizes observable functions, is learned through a least square loss function encoding linearity of Koopman with A and B fixed. The results indicate the deep Koopman method outperforms the extended DMD (EDMD), which selects numerous observation functions manually, in terms of state prediction accuracy and transient stabilization [41]. [42] devises a deep input-Koopman learning approach to capture the nonlinear suite of frequency regulation dynamics, including open-loop swing dynamics, market behavior, and spoofing mechanism, thereby enabling a controller synthesis to be robust to price spoofing attacks. DNN is leveraged to characterize the infinite-dimensional dictionaries of observables defined on the state and the input, respectively, by enforcing the linear evolution law of Koopman. A similar scheme is adopted in [44] to learn the transient dynamics of a PS.

The above results and analysis exemplify that the PINN technique can extract key representations from the nonlinear state-space model of complex systems or approximate the nonlinear systems directly. Either way, PINNs can constitute control synthesis in conjunction with controllers to facilitate better stability and resilience of the system. Another primary use of PINN in the power system control domain is the design of data-driven controllers, especially those based on reinforcement learning (RL). RL, as a computational paradigm for multi-stage decision optimal control, is widely used in optimal control of PSs, such as frequency regulation (FR) [81], automatic generation control (AGC) [46], load-frequency control (LFC) [47], etc. RL is a trial-and-error approach in which the decision-making agent takes actions based on observed states and updates its policies based on rewards received from the environment with the evolvement of value functions. RL algorithms can be divided into model-based RL and model-free RL, where the former establishes a predictive model of the environment while the latter does not. The breakthroughs in DL gave rise to deep reinforcement learning (DRL), which uses DNNs to approximate the value functions, the policies, and even the environment itself. A survey paper [82] and a vision paper [83] comprehensively reviewed and projected the RL/DRL-based control on PSs, respectively. Considering that optimal control is a general concept under the context of PSs, we divide our discussion of control-related problems into two major parts. This subsection reviews and discusses the junction of the PINN and RL/DRL on dynamics control of PSs. The RL/DRL on steady/quasi-steady optimal control problems that are usually formulated as optimization problems and algebraic equations will be covered in Sec. III-D. The environments of DRL are different under these two conditions (different time scales and different numerical models), whereas they have intense cross-fertilization with each other in terms of design ideas.

Work presented in [43] deals with FR using an actor-critic approach, which is a typical class of DRL methods. To ameliorate the long-term system performance, a NN is utilized to approximate the strategic utility function (SUF), defined as the weighted sum of performance index over a rolling horizon and designed to impose a penalty on the magnitudes of the state (frequency deviation and regular control-related signals) and control output (adjustable active power of distributed energy resources). SUF originates from the physical phenomenon that power devices cannot be operated under a significant frequency deviation for a long time. The application of NN approximation and its combination with DRL makes SUF trainable and enables SUF to aid the update of the optimal controller. [46] demonstrates the utilization of imitation learning and transfer learning to accelerate the RL agent, which is devised to conduct the AGC of a virtual power plant. The proposed imitation learning technique collects effective state action pairs from swarm intelligence algorithms to efficiently train the RL agent's value function to improve the dynamic performance of AGC. [47] introduces the emotional learning technique to the RL agent on LFC, where action, learning rate, and reward function are adjusted through a two-layer artificial emotion quantizer, thereby making the controller adapt to various load disturbances. [45] proposes a global-local hybrid control architecture to damp inter-area oscillations, where local controllers are power system stabilizers (PSSs), and a heuristic dynamic programming-based DRL agent is adopted as the global controller. A value priority strategy inspired by the Lyapunov energy function coordinates the local and global controllers to achieve coherent damping.

2) Summary and Potential Research Directions: Established model-based control techniques, such as MPC-based and rule-based control strategies, usually involve the complex physical model of the system characterized by a set of ordinary or partial differential equations, which compromises the tractability and scalability. In contrast, pure data-driven control techniques circumvent the shortcomings above by avoiding using costly numerical methods, whereas these techniques suffer poor extrapolability and insufficient safety guarantees. As an intermediate technique, PINN can get the best of both sides. PINNs can assist in resolving control problems of PSs through multiple forms. First, PINNs can be harnessed to learn the physically relevant control-oriented state-space model of real-world systems within the design of control setups [84], [85]. Incorporating these PINNs-based state-space models into the control synthesis is critical for the PINNs to be extended to control-related problems. Second, considering that the deep Koopman operator can aid the optimum control, developing a proper control schema to take advantage of this powerful tool becomes necessary [41].

The fusion of the first principles, domain knowledge, control practice, and DRL-based control is another promising research direction, of which the crux is related to the core components of DRL and can be summarized as follows. 1) *Reward design*. In addition to simply designing reward functions based on physical information, preference-based RL algorithms are worth consideration in the domain of PSs because they can accommodate a broader spectrum of tasks and integrate the domain knowledge of experts [86]. The conventional RL paradigm suffers from manually curated numeric reward shaping. In contrast, preference-based RL learns such a reward through a preference-based feedback signal that reflects relative instead of absolute utility values and is not subject to a predefined reward landscape. 2) *Environment*. The environment in DRL is often a real-world system or a model-based simulator, which PINN-based surrogate model can replace to enrich the training set and accelerate the training. Besides, Bayesian RL is worth further studies in light of the fact that prior knowledge can be integrated into the parameters of the Markov model in model-based RL using this technique [87]. 3) *Environment/Agent interaction mode*. Both the coordination or competition mechanism of agents in the multi-agent DRL (MADRL) setting and surrogate model/agent interaction can be ameliorated by incorporating physical knowledge. Moreover, [45] exemplifies the cooperation of the DRL controller and model-based controller. 4) *Physics-informed value function and policy design*. Bayesian RL provides an elegant approach to convey physical priors via value function or policy class in model-free RL [87]. The work in [88] replaces the original action-value functions parameterized by DNNs with nonlinear MPC schemes, of which the closed-loop behavior is well-studied in control theory, endowing the RL controller with better reliability and interpretability. Such practices in the control community can be considered in PSs as well. 5) *Initialization of policy from physical guidance*. An example in this aspect is [46]. In terms of the further development of physics-informed DRL-based control, it is advised to implement the safe exploration of DRL with the guidance from domain knowledge and practical constraints of sensors and actuators, thereby certifying the safety of a learned control policy and its closed-loop behavior [89]. One way to achieve this is embedding priors into constrained Markov decision processes using Bayesian approaches. Another non-negligible issue is the scalability of DRL, which limits the direct application of DRL to large-scale PSs, considering the combinatorial explosion of state and control vector. Integrating PINNs with the MADRL framework is a promising direction for solving the scalability problem.

The progress of PINNs in solving PDEs [12] unlocks opportunities in capturing the dynamics of PSs with the reduced computational burden. This is due to the automatic differentiation feature of the current DNN framework, which fits well with the numerical characteristics of PDEs. Compared with the traditional numerical methods, PINNs are more efficient in the test phase. Compared with the pure data-driven DL method, PINNs need less training data and reduce the burden of training data acquisition. This is valuable in dynamic analysis because the training data of DL methods often come from time-consuming simulation experiments or rare events in the real world. The existing literature usually uses one or several dynamic processes to verify the performance of PINNs. A statistical analysis on a broader range of dynamic processes can render the performance of PINNs more convincing.

In general, most developments in solving PDEs using PINN can be used for the application of dynamic analysis of PSs [90]. This research direction has many questions in need of further investigation. 1) In these applications, the training of PINN is dependent on the availability of collocation points. How to select representative collocation points becomes crucial. 2) The current literature mainly reports the results of PINN on small-scale systems, e.g., SMIBS [13] and Kundur two-area system [37]. Therefore, it is very meaningful to investigate the scalability of PINN to expand the scope of PINN applications. 3) Dynamic problems of PSs are strongly related to the stability and security of the system, whereas the derivation of DNNs is based on probability theory and statistics. Existing research recognizes the critical role played by performance guarantees of NNs [91], [92]. [91] investigates the verification of NN behavior and develops a verification framework in PSs applications, including security assessment and small-signal stability. The verification procedure is dependent on the reformulation trick of the ReLU, with which the MILP is formulated to determine the safe margin of a specific operating point. The same techniques can be transplanted to PINN, which can enhance the interpretability of PINN and give more confidence to the users of the PINN on PSs. [92] explores error bounds for the ML algorithm on the regression problems in PSs and investigates the relationship between the amount of physical knowledge and the bounds. The above references can inspire the development of PINN on performance guarantee. 4) How to incorporate discrete events (e.g., protection actions) into PINN, and how to perform PINN on small-signal stability and voltage stability are yet to be studied.

C. Power Flow Calculation

1) Literature Review: Power flow (PF) calculation is the basis of the steady-state analysis of PSs. Conventional approaches employ numerical methods, such as the well-known Newton-Raphson method and Gauss-Seidel method, to conduct PF calculation, which is dependent on the availability of PFEs. Traditional methods can not adapt to modern power systems under certain scenarios because of the difficulty of maintaining the PFEs. For example, it is difficult to obtain the topology information of active distribution networks embedded with autonomous microgrids in real-time. To address these issues, various data-driven approaches, i.e., LR [93] and SVR [94] have been adopted to solve the PF problem. However, they still suffer from either over-simplicity or scalability issues.

To resolve these issues, [48] designs a semi-supervised auto-encoder architecture with the framework of multi-task learning. The learning goal is two-fold: 1) minimizing the voltage prediction error on the level of encoder only, which features the regular supervised learning; 2) minimizing the power mismatches on the level of encoder-decoder, which features the unsupervised learning. The encoder part is a multilayer perceptron (MLP) that serves as a PF solver by learning the mapping from PF inputs to nodal voltages. The decoder part carries out the auxiliary task, which needs to ensure the NN can reconstruct the power injections with the estimated voltages. The weights of the encoder can be updated along with the training of the decoder. To shrink parameter search space and improve generalization, the infusion of PFEs is reflected in the design of the decoder. A bilinear NN is employed to imitate the nodal admittance matrices. What is more, given the topology of the PS is available, the parameters of the decoder can be pruned by encoding the available adjacent matrix as a hard attention mask. The tests on IEEE 118 bus system show that the PINN can perform stably with the increase of the level of noisy outliers, as opposed to LR and SVR.

The advances of GNNs bring fresh perspectives on solving PF. [49] proposes a purely data-driven approach based on graph convolutional networks (GCNs) and reports promising results on real-world power grids such as Texas (544 generators and 2345 lines) and East Coast (835 generators and 5411 lines) systems. [50] employs an encoder/decoder framework and a messages propagation mechanism among neighboring nodes to solve PF. There are three major limitations in [50], i.e., strong physical assumption, dependency on prepossessing, and dependency on the Newton-Raphson solver to generate a training set. As responses to the above limitations, [51] proposes an independent GNN PF solver which minimizes the weighted average of violation of Kirchhoff's laws iteratively rather than mimics the behavior of another solver. The proposed GNN PF solver utilizes the hybrid physics-DL paradigm, which associates the learnable NN parts with non-learnable physics-based parts. The outputs of physics-based parts, which are global active compensation and local power imbalance, are the components of the loss function to the learnable NN. They are based on the PFEs and used to tune the behavior of GNN. The experiment results show that the GNN PF solver is superior to the direct current PF (DCPF) in efficiency and accuracy over IEEE 9, 14, 30, and 118 bus systems. Another benefit of the proposed solver is that the learned weights are generalizable to different grids.

2) Summary and Potential Research Directions: As an alternative to alternating current PF (ACPF) and DCPF, the further expansion of the application of the PINN in PF calculation is facing a dilemma of accuracy and efficiency. On the one hand, the users of ACPF prefer theoretically solid solutions because they are concerned about the realistic operating states of PSs. On the other hand, under the scenarios where the training time is nonnegligible (e.g., real-time training is required), PINN has no significant advantage over DCPF, for which a fast and reliable solution (by solving a system of linear equations) is already available.

The preliminary studies of applying PINN to PF calculation open the doors for solving PF problems with a more intricate formulation. 1) Similar to the SE problem, the three-phase unbalanced PF formulation is not explored under the context of PINN. 2) Moreover, current PINN methods cannot encode the discrete events in PF calculation, e.g., reactive power limits for generators, because such events would cause non-differentiable NN formulation. 3) Another topic relevant to the PF problem is voltage stability. Previous studies on voltage stability are conducted by analyzing the singularity of the Jacobian matrix. It would be appealing to construct the multi-task PINN-based PF solver, which outputs the voltage phasors and voltage stability margin concurrently.

D. Optimal Power Flow

1) Literature Review: Different from the PF problem, the Optimal Power Flow (OPF) problem is an optimization problem. With the objective of minimizing cost, maximizing security margin, minimizing pollution, etc., OPF decides the voltage set-point and power output of generators that meet the load demands and satisfy the practical constraints of PSs. However, traditional optimization-based OPF faces one dilemma, where the execution frequency and computational complexity should be balanced.

To overcome this challenge, [52] approaches the OPF via a PINN. Specifically, it reformulates the OPF as the Lagrangian dual of an empirical risk minimization problem, where the risk is defined as the violation degree of constraints. Stacked MLPs with skip-connection serve as the OPF solver, producing a high-quality solution. The physical and engineering constraints under the Lagrangian framework are integrated into the DNN via the violation term in the loss function. The decoder-encoder framework with four subnetworks is used for four different target variables (i.e., voltage magnitudes, phase angles, and active and reactive power generations), respectively. Compared to the ubiquitous direct current OPF (DCOPF) model, the computational results over several test systems show superior accuracy and efficiency of PINN. A solution based on DRL is proposed to resolve the real-time OPF in [58]. A modified DRL algorithm is examined, featuring the augmented objective action-value function and the elimination of the critic network. [58] designs an action-value function to imitate Lagrangian function and derives the gradient ascent step analytically for updating actor-network without the help of a critic network.

To deal with the challenge of frequent model changes of OPF, which hinders the direct application of DNNs, sample-efficient learning models that generalize well are motivated. One way to achieve that is by prudently designing the DNN architecture using prior information. [53] develops objective functions based on Karush-Kuhn-Tucker (KKT) conditions to train an input-convex DNN, featuring ReLU function and nonnegative weights. [54] puts forth a novel approach for training DNNs to predict OPF solutions by matching not only the OPF minimizers but also their sensitivities (partial derivatives collected in a Jacobian matrix) concerning the OPF parameters, such as load demands. The calculation of sensitivity matrices depends on the DL packages' automatic differentiation feature. Extensive numerical tests on three benchmark power systems have demonstrated that with a modest increase in training time, sensitivity informed-DNNs attain the same prediction performance as conventionally trained DNNs by using roughly only 1/10 to 1/4 of the training data.

OPFs are often used in security-critical PS applications where the operator is particularly interested in the robustness of the scheduling results given by the OPF, i.e., the performance under worst-case scenarios. Based on this challenge, [55] tries to integrate the KKT conditions of DCOPF into the loss function of training NN. Then, by restructuring NN with the ReLU activation function as MILP problem and solving it, [55] supplies worst-case guarantees for constraint violations and optimality. Extending the work on DCOPF [55], [56] addresses the worst-case guarantees problem in alternating current OPF (ACOPF). ACOPF is formulated in the Cartesian coordinates, and KKT conditions are derived based on the Lagrangian function of ACOPF. A multi-task NN framework is developed to simultaneously predict power generation setpoints, voltage setpoints, and dual variables. A PINN layer that hardcodes KKT conditions is added to the NN, and the outputs of this layer represent the disparities in the KKT conditions. Collocation points collected from random input values are used as the extended training data for better performance of PINN. [56] uses MILP and Mixed Integer Quadratic Constrained Quadratic Programming (MIQCQP) to determine the worst generation and line flow constraint violations. Another interesting insight from [56] is that it is prohibitively expensive to determine the worst line flow constraint violations under certain cases due to the intractable MIQCQP.

Thanks to the excellent function approximation capability, DNNs serve as the inverter policy in [57], which receives grid scenarios and then determines the reactive power set-point of each inverter. In the light of the distributed operation characteristics of the inverters, a two-tier DNN architecture (top tier for offline training and bottom tier for real-time operation) is proposed to comply with the physical communication architecture. The policy DNN also features the decoupled structure, encoding the distributed communication architecture of smart inverters.

DRL has also been closely related to optimal steady/quasi-steady optimal control problems [95]. A growing body of literature recognizes the superiority of the physics-informed model-free DRL on voltage/reactive (Volt/Var) power control. Model-free DRL refers to learning optimal policies directly. Inspired by [58], [59] devises a two-stage MADRL framework for Volt/Var control, in which the action-value function comprises power loss, voltage violation, and reactive power capability violation. The usage of voltage sensitivity information guides the derivation of the gradient ascent step and excludes the pure data-driven critic network. The MADRL framework in [59] has the issue of scalability with the increase of photovoltaic (PV) inverters, and [60] mitigates this issue by using Volt/Var sensitivity to partition the system to several sub-areas and developing the DRL agents on the sub-area level. Moreover, considering the availability of the accurate distribution network model and parameters, a Gaussian Process Regression-based surrogate model is leveraged to mimic the response of the distribution network under varied operating status and substitute the load flow-based environment setting. Based on [60], [61] employs the sparse pseudo-Gaussian process as a surrogate model to achieve few-shot learning.

Another major application branch of PINN in PSs is imitation learning (IL). Many IL-related studies stem from the field of DRL, which is a tool for decision-making using NNs. Instead of training NNs from scratch, IL emphasizes using prior knowledge or expert demonstration to give DNN a better start point, which coincides with the idea of using physical prior knowledge to initialize the weights of PINN. [62] regards the restoration strategy of PSs based on integer programming as the teacher of DNN and trains DNN via supervised learning. The DRL training process in power grid control is scrutinized in [63], which discovers the DRL agent learns mainly from the last successful steps. Subsequently, a sophisticated sample selection mechanism is proposed to help build the training set of IL, which supplies the DRL agent with a decent initialization policy. Both [62] and [63] report that physics-informed IL can boost the training efficiency of DRL agents applied in PSs.

2) Summary and Potential Research Directions: In contrast to PF, the application of PINN in OPF is more active. Because optimization problems are often considered to be computationally intensive and suboptimal solution tolerant, PINN has significant potential to improve the solution efficiency. The current study of PINN focuses on integrating the KKT condition and convexity of the formulation. It is also a common strategy to put the constraint violation of OPF into the loss function of the PINN through the Lagrangian method. The current research mainly focuses on standard ACOPF or DCOPF formulation considering line power flow constraints and PFEs. As an extension of OPF, the unit commitment problem and its constraints on ramping and minimum up/down time constraints have not been explored. Furthermore, the summary and suggestions on physics-informed DRL-based control presented in Section III-B are still effective under the context of steady/quasi-steady optimal control problems. A prominent issue in this domain is that the action frequencies and response speeds of control devices in some cases are different, such as the inverters and on-load tap changers in voltage control, which necessitates a coordination mechanism to make the control scheme more practical. It is expected that the usage of PINN will show tremendous potential in model-free RL. The reason is that the integration of physical knowledge is a solution to the issues of low data efficiency and bootstrapping bias, which are common deficiencies of model-free RL, by producing high quality and physics-consistent samples.

The potential directions of future work in this niche can be summarized as follows. 1) Both PF and OPF are based on PFEs. Therefore, generic first principles based on PFEs can be used to further guide PINN, e.g., the symmetric structure of the admittance matrix. In a high voltage transmission network, the reactance of line parameters is much greater than the resistance. Active power and reactive power mainly affect the voltage phase angle and magnitude, respectively. For light-loaded power lines, the phase angles at both ends of the line have little difference. The element corresponding to the nodal reactive power in the admittance matrix is usually much smaller than the nodal self-admittance [96]. 2) In addition to linearity, other convexities in OPF research may be integrated into PINN. For example, the second-order cone relaxation is commonly used in distribution network OPF [97]. 3) Despite the great advancements in [55], [56], further research should be undertaken to explore how to derive the worst-case guarantees for complicated constraints of ACOPF, e.g., voltage stability constraints, security constraints under contingency conditions, time-coupled constraints from ESSs, etc.

E. Anomaly Detection and Location

1) Literature Review: In order to avert power outages and secure stable operation, it is of great importance to detect the faults of PSs as early as possible. Detecting high impedance faults (HIFs) for protection relays remains a challenging task because of its insignificant over-current phenomenon [64]. The existing data-driven HIF detection methods are sensitive to noise and low harmonics. The requirement of a sufficient number of labeled faulty samples is another hurdle. To resolve these issues, a physics-informed convolutional auto-encoder (PICAE) is proposed in [64] to detect HIFs without labeled data for training. [64] designs an unsupervised convolutional AE architecture, of which the input and output are measured and reconstructed voltage phasors, respectively. Then, the well-trained AE can output the reconstructed voltage phasor that is very close to the input measured voltage phasor. If the error between the input and output of AE exceeds the preset threshold, this group of measured voltage phasors will be given a lower confidence score, which is determined as the faulty voltage. Furthermore, given that the HIF features the approximate elliptical trajectory of voltages and current, [64] incorporates the elliptical regularization term in the loss function to train the convolutional

AE, thereby regulating the weight update of both the encoder and decoder simultaneously. The comparative studies between PICAE, AE, principle component analysis (PCA), and ellipse regression (ER) are conducted on the IEEE 34-node test feeder. The result substantiates that the elliptical regularization, i.e., hidden physics law, endows the PICAE with robust detection performance towards HIFs. Moreover, proper PMU placement can mitigate the negative impact of partial measurements.

Fault location is the prior step in restoring power to the PSs after experiencing an outage, which is closely tied to the reliability and resilience of the system. GCNs are exploited in [65] to enhance the accuracy of fault location by encoding the spatial correlations of the buses. Taking a step further, [66] extends the scope of fault location to the scenarios of low observability, labeled data deficiency, and various data distributions. A two-stage scheme is adopted in [66], where stage I informs the PINN of the topological adjacency matrix based on the K nearest neighbors approach and Gaussian kernel. Considering the localized characteristics of the fault, Stage II performs the label propagation through graph convolutional layers, which can overcome the issue of labeled data deficiency. As opposed to baselines without physical prior, PINN can maintain satisfactory accuracy (95%-100%) during the plummet of label rate, load variations, or topology changes.

2) Summary and Potential Research Directions: So far, the applications of PINN on anomaly detection and location are mostly problem-dependent. One example is the approximate elliptical relationship of voltages and current in the HIFs [64]. However, this kind of engineering practice is not universal to all types of faults, though the way of integration of a priori knowledge, e.g., loss function design, is more generic. The primary motivation of PINNs on the anomaly detection and location is that the acquisition cost of abnormal data in PSs is very high, while training DNN often requires substantial data. By fusing the first principles, PINN significantly compresses the search space of network parameters and reduces the requirements of large quantities of training data.

Potential directions of future work can be summarized as follows. 1) In the procedure of abnormal signal processing, the physics-based kernel is an excellent substitute for the data-driven kernel because it can alleviate the over-fitting issue. For example, in [98], a delicate convolution layer is designed to depict the similarity of the input signal (data) and the fault feature signal (physics) from spectral kurtosis analysis and amplitude demodulation; 2) The transfer learning of PINN should be studied to build a bridge between the detection of common anomalies or easily simulated anomalies and the detection of rare anomalies, thereby improving the sample efficiency further [99].

F. Model and Data Synthesis

1) Literature Review: The synthetic power system model is a new and active research direction of PINNs. There are two fundamental problems on this research topic: how to make DNN understand unstructured power grid topology and data and how to make DNN generate a physically feasible power grid model. [67] tries to map heterogeneous distribution feeder circuit data (e.g., device connectivity, phase information, length, and current ratings) to the adjacency matrix and attribute matrix, respectively. Wasserstein generative adversarial networks (GAN) are employed to generate the synthetic feeders, where GCN is adopted in discriminator to distinguish the real and generated feeders. [68] develops a graph distribution learning method to synthesize large-scale transmission networks based on the physical characteristics of the power grid. [68] formulates the PS synthesis as a probabilistic learning problem of the probability distribution functions of the buses and lines. The modularity of PSs, which is considered as the topological prior of PSs, is exploited to divide the realistic model into communities, followed by modeling independent nodal and edge features by RNN in each community. The synthetic models from PINN show more realistic topological properties and power flow statistics relative to Gaussian Mixture Model (GMM).

2) Summary and Potential Research Directions: GAN has been an influential generative framework in DL, and it is used extensively in model and data synthesis. The existing related research lacks a discussion on the follow-up application of the synthesized model and data. It would be beneficial to use synthetic data or models in realistic applications to substantiate the practicability.

Potential research gaps include: 1) The current research focuses on the steady-state model. The application of generative models may be extended to generate dynamic models with smaller time scales, e.g., generating the parameters of the third-order model of synchronous generator. 2) Besides synthesizing the snapshot of the model, generative networks and LSTM can jointly characterize the time correlation and generate the synthetic systems evolving within a period, similar to generating lyrics and videos[100].

G. Miscellaneous Applications

Ref. [69] exhibits the interesting results on predicting the power of wind farms using physics-informed GNN. The wind farm is modeled as a graph, and specialized message-passing GNN with physics prior serves as the predictor. The architecture of GNN is explicitly designed for imitating the wake-deficit model, which is a commonly used numerical model considering the influence of upstream wind turbines. Physical knowledge dominates the design of NN architecture in [69].

Except for PINN, the integration of physical knowledge also promotes the application of ML algorithms and sampling techniques in PSs. Compared with Monte Carlo sampling, Importance Sampling (IS) techniques are more efficient and effective in estimating the probability of overload of lines [101]. [101] attempts to improve the IS techniques by integrating the distilled physical knowledge from DCPF and Gaussian power injections. [102] and [103] report the applications of physics-informed Gaussian process on probabilistic DSE [102] and frequency oscillations [103], respectively. The dynamic characteristics of PSs are extracted to guide the selection of covariance and kernel function in [103]. The unknown mechanical wind power in the swing equations is formulated as a random process in [102], where the Monte Carlo approach deals with

TABLE II
TABLE OF LITERATURE CLASSIFIED BY THE PROPERTIES OF APPLICATIONS

Property	Category	Reference
Forward or inverse	Forward	[13], [25]–[29], [31]–[34], [36]–[39] [41]–[57], [59], [60] [61]–[64], [64]–[66], [69]
	Inverse	[13], [30]–[32], [34], [48] [35], [67], [68]
Steady-state or dynamic	Steady-state	[25]–[33], [48]–[55] [56], [57], [59]–[63], [65], [67], [68]
	Dynamic	[13], [34]–[39], [41], [42], [44], [64], [66] [43], [45]–[47]
Operation or control	Operation	[13], [25]–[34], [36]–[39], [44], [48]–[51] [35], [52], [53], [55], [56], [62], [64]–[66]
	Control	[41], [42], [54], [57], [59]–[61], [63] [43], [45]–[47]

estimating the statistics of PSs. [104] reports the advance of a physics-informed extreme learning machine (ELM) on the OPF of PSs. ELM is famous for its limited learning capacity and fast training efficiency. In order to cater for the nonlinear and complicated OPF problem, [104] proposes a three-stage ELM learning framework, where the line flow equations and nodal PFEs are decomposed and reflected in ELM in different stages.

H. Classification

In addition to the classification according to the applications of PSs, Table II classifies the collected references into different groups, i.e., in terms of forward vs. inverse problems; power system operation vs. control; and steady-state vs. dynamic models. The inverse problems here refer to the process of relying on the results to infer the causes (e.g., identification of the parameters of a dynamic system using the observables), and the forward problems are just the opposite.

I. Discussion

The aforementioned references and corresponding analysis exhibit the advantages of PINNs over pure data-driven and first-principle-based models (e.g., accuracy, generalizability, efficiency, interpretability, training data requirement, etc.) in quantitative and qualitative forms within the context of PSs. Admittedly, the pure data-driven method may perform better in some specific applications. For example, compared to AlphaGo, AlphaGo zero performs better without the aid of human experience [105]. However, it should be noted that the rules and boundaries of go are encoded in the training of AlphaGo zero. Hence, AlphaGo zero is not completely prior-knowledge-agnostic. In addition, the improvement of AlphaGo zero is primarily due to the increase of computing resources and computing time. Therefore, PINN is a promising choice for tasks with partial rules and unknown boundaries in PSs, especially when computing resources are limited and labeled data is scarce. In the practical application of PINNs, the integration of physical knowledge depends on the availability of physical knowledge. For example, in the PINN for power flow calculation in [48], three decoder architectures mapping voltage to node power injection are devised, which correspond to the low to the high availability of PS information. The first architecture is the decoupled MLP, which is a data-driven scheme. The second architecture is bilinear NN, which harnesses the structural property of ACPFEs. The third architecture is topology-aware bilinear NN, which assumes that the topology of the PS is known and is used to zero out the corresponding neurons in bilinear NNs. Users are advised to select the appropriate PINN framework according to the application scenario.

IV. LIMITATIONS AND PATH FORWARD

The diversified successful applications of PINN unveil the potential of the PINN from multiple perspectives. This section aims to provide a broad and general vision for the future development of PINN on the open problems in PSs, along with pointing out current limitations.

The path forward for PINNs on PSs should align with the modern and future development of PSs. In terms of application scenarios, harnessing PINNs to accommodate increasing integration of distributed energy resources, electric vehicles, microgrids, power electronics devices, advanced communication, metering, and edge computing infrastructure, etc, is of particular interest. It is noticed that although Section III-D reports the extensive applications of PINN on OPF, as a variant of the OPF problem, the PS planning problem has not been solved via PINN. The challenge here is that the planning problem is a mix-integer problem, and current PINNs cannot handle the hybrid output that contains both discrete and continuous variables. The advance in solving mixed-integer programs using NNs [106] may pave the path for addressing this challenge. Furthermore, numerous signals or data in PSs come in the form of complex numbers, such as voltage, power, impedance, etc. Complex-valued PINN is one of the potential directions that can tailor the PINN for applications in PSs better.

Performance guarantee and verification are the keys to further recognition of PINN in power engineering. Through providing quantitative evaluation criteria to users, it can also stimulate the application of PINNs in real PSs. Current relevant research focuses on the reformulation trick of the ReLU and MILP. To accommodate more security-critical challenges in PSs, it is paramount to develop the robustness certification of PINN to various input data distributions, network architectures, layers, and activation functions [107].

Transfer learning is deeply valued in the DL community because it can enhance the sample efficiency and training efficiency significantly, especially for problems with a small amount of training data. The adaptability of PINN has rarely been discussed in the existing literature. Here are some power system scenarios that may greatly benefit from the utilization of PINNs with transfer learning: 1) PINNs across different test systems; 2) PINNs in consideration of network expansion and reconfiguration; and 3) PINN on SE tasks across multiple levels of observability.

At present, we have observed that PINNs are applied in a variety of DNN architectures, such as MLP, RNN, and convolutional NNs. As one of the most popular DL architectures, the attention mechanism has been very successful in natural language processing [108] and has been transplanted to PSs recently [109]. It will be exciting to integrate the first principles into the attention mechanism. Quantum computing (QC), which depends on quantum entanglement and quantum superposition, is expected to achieve overwhelming superiority in computational efficiency in some specific applications. As a cutting-edge research area, quantum ML explores the interplay of ideas from QC and ML and extends the implementation of ML algorithms on a universal or near-term quantum computer. As a form of hybrid physics-DL models, the differentiable quantum circuits can be devised to mimic and replace some DL components, thereby tackling problems in the power and energy industry. Preliminary explorations include quantum transfer learning [110], quantum GAN [111], and quantum graph RNN [112], etc. Besides, DNNs can also be used to distill key factors from experimental data to guide the discovery of unknown governing equations [113], [114]. Hence, it is promising that symbolic regression [115] in conjunction with PINN can serve to discover the underlying dynamics of PSs. Last but not least, analogous to the ImageNet [116] in computer vision, Gym in DRL [117], and Power Grid Lib in OPF [118], benchmark libraries are needed to evaluate the emerging PINN algorithms.

V. Conclusion

This survey provides a comprehensive review of PINN on PSs. Revolving around the ingredients of NN, four common PINN paradigms are presented, complemented by easy-to-understand practical examples. PINN and its applications in the domain of state/parameter estimation, dynamic analysis, power flow calculation, optimal power flow, anomaly detection and location, model and data synthesis, etc., are reviewed. We provide a comprehensive summary of the specific application, integration mode of physical prior, baseline methods, and qualitative and quantitative effects of PINN on PSs. Combining with the characteristics of PSs and the state-of-the-art DL techniques, this survey sheds light on the potential research directions of PINN on PSs.

Acknowledgment

The views expressed herein do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

[1] A. Karpatne *et al.*, "Theory-guided data science: A new paradigm for scientific discovery from data," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 10, pp. 2318–2331, Oct. 2017.

[2] S. Zhang, M. Wang, S. Liu, P.-Y. Chen, and J. Xiong, "Fast learning of graph neural networks with guaranteed generalizability: One-hidden-layer case," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 11268–11277.

[3] J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, "Integrating scientific knowledge with machine learning for engineering and environmental systems," *ACM Comput. Surv.*, Jan. 2022, doi: 10.1145/3514228.

[4] X. Jia *et al.*, "Physics-guided machine learning for scientific discovery: An application in simulating lake temperature profiles," *ACM/IMS Trans. Data Sci.*, vol. 2, no. 3, pp. 1–26, 2021.

[5] K. Schütt, P. Kindermans, H. Sauceda, S. Chmiela, A. Tkatchenko, and K. Muller, "Schnet: A continuous-filter convolutional neural network for modeling quantum interactions," *Adv. Neural Inf. Process. Syst.*, vol. 2017, pp. 992–1002, 2017.

[6] G. R. Schleder, A. C. Padilha, C. M. Acosta, M. Costa, and A. Fazzio, "From dft to machine learning: Recent approaches to materials science-a review," *J. Phys.: Mater.*, vol. 2, no. 3, 2019, Art. no. 032001.

[7] M. Reichstein *et al.*, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.

[8] D. Xiao *et al.*, "A reduced order model for turbulent flows in the urban environment using machine learning," *Building Environ.*, vol. 148, pp. 323–337, 2019.

[9] M. Khodayar, G. Liu, J. Wang, and M. E. Khodayar, "Deep learning in power systems research: A review," *CSEE J. Power Energy Syst.*, vol. 7, no. 2, pp. 209–220, 2021.

[10] F. Noé, A. Tkatchenko, K.-R. Müller, and C. Clementi, "Machine learning for molecular simulation," *Annu. Rev. Phys. Chem.*, vol. 71, pp. 361–390, 2020.

[11] R. Rai and C. K. Sahu, "Driven by data or derived through physics? a review of hybrid physics guided machine learning techniques with cyber-physical system (CPS) focus," *IEEE Access*, vol. 8, pp. 71050–71073, 2020.

[12] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019.

[13] G. S. Misyris, A. Venzke, and S. Chatzivasileiadis, "Physics-informed neural networks for power systems," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2020, pp. 1–5.

[14] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1. Cambridge, MA, USA: MIT Press, 2016.

[15] S. Shah, D. Dey, C. Lovett, and A. Kapoor, "Airsim: High-fidelity visual and physical simulation for autonomous vehicles," in *Field and Service Robotics*. Berlin, Germany: Springer, 2018, pp. 621–635.

[16] A. Daw, R. Q. Thomas, C. C. Carey, J. S. Read, A. P. Appling, and A. Karpatne, "Physics-guided architecture (PGA) of neural networks for quantifying uncertainty in lake temperature modeling," in *Proc. SIAM Int. Conf. Data Mining*, 2020, pp. 532–540.

[17] J. Sun, Z. Niu, K. A. Innanen, J. Li, and D. O. Trad, "A theory-guided deep-learning formulation and optimization of seismic waveform inversion," *Geophysics*, vol. 85, no. 2, pp. R87–R99, 2020.

[18] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, "A review of graph neural networks and their applications in power systems," *J. Modern Power Syst. Clean Energy*, vol. 10, no. 2, pp. 345–360, Mar. 2022, doi: 10.35833/MPCE.2021.000058.

[19] A. Samadi, L. Sóder, E. Shayesteh, and R. Eriksson, "Static equivalent of distribution grids with high penetration of PV systems," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1763–1774, Jul. 2015.

[20] S. A. Billings, *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. Hoboken, NJ, USA: Wiley, 2013.

[21] Y. Chen, S. M. Mazhari, C. Chung, S. O. Faried, and B. C. Pal, "Rotor angle stability prediction of power systems with high wind power penetration using a stability index vector," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4632–4643, Nov. 2020.

[22] E. J. Parish and K. Duraisamy, "A paradigm for data-driven predictive modeling using field inversion and machine learning," *J. Comput. Phys.*, vol. 305, pp. 758–774, 2016.

[23] K. Yao, J. E. Herr, D. W. Toth, R. Mckintyre, and J. Parkhill, "The tensormol-0.1 model chemistry: A neural network augmented with long-range physics," *Chem. Sci.*, vol. 9, no. 8, pp. 2261–2269, 2018.

[24] C. Ruben *et al.*, "Hybrid data-driven physics model-based framework for enhanced cyber-physical smart grid security," *IET Smart Grid*, vol. 3, no. 4, pp. 445–453, 2020.

[25] L. Zhang, G. Wang, and G. B. Giannakis, "Real-time power system state estimation and forecasting via deep unrolled neural networks," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 4069–4077, Aug. 2019.

[26] Q. Yang, A. Sadeghi, and G. Wang, "Data-driven priors for robust PSSE via gauss-newton unrolled neural networks," *IEEE J. Emerg. Select. Top. Circuits Syst.*, vol. 12, no. 1, pp. 172–181, Mar. 2022, doi: 10.1109/JETCAS.2022.3142051.

[27] J. Ostrometzky, K. Berestizshevsky, A. Bernstein, and G. Zussman, "Physics-informed deep neural network method for limited observability state estimation," in *Proc. Workshop Auton. Energy Syst. NREL Golden CO 8/19-20/2020,* 2020, pp. 1–6.

[28] A. S. Zamzam and N. D. Sidiropoulos, "Physics-aware neural networks for distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4347–4356, Nov. 2020.

[29] M.-Q. Tran, A. S. Zamzam, and P. H. Nguyen, "Enhancement of distribution system state estimation using pruned physics-aware neural networks," in *Proc. IEEE Madrid PowerTech*, 2021, pp. 1–5.

[30] W. Wang and N. Yu, "Estimate three-phase distribution line parameters with physics-informed graphical learning method," *IEEE Trans. Power Syst.*, early access, 2021, doi: 10.1109/TPWRS.2021.3134952.

[31] L. Pagnier and M. Chertkov, "Physics-informed graphical neural network for parameter & state estimations in power systems," 2021, *arXiv:2102.06349*.
[32] L. Pagnier and M. Chertkov, "Embedding power flow into machine learning for parameter and state estimation," 2021, *arXiv:2103.14251*.
[33] L. Wang, Q. Zhou, and S. Jin, "Physics-guided deep learning for power system state estimation," *J. Modern Power Syst. Clean Energy*, vol. 8, no. 4, pp. 607–615, 2020.
[34] J. Stiasny, G. S. Misyris, and S. Chatzivasileiadis, "Physics-informed neural networks for non-linear system identification for power system dynamics," in *Proc. IEEE Madrid PowerTech*, 2021, pp. 1–6.
[35] G. Tian, Q. Zhou, R. Birari, J. Qi, and Z. Qu, "A hybrid-learning algorithm for online dynamic state estimation in multimachine power systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 12, pp. 5497–5508, Dec. 2020.
[36] J. Stiasny, S. Chevalier, and S. Chatzivasileiadis, "Learning without data: Physics-informed neural networks for fast time-domain simulation," in *Proc. IEEE Int. Conf. Commun., Cont., Comput. Technol. Smart Grids (SmartGridComm)*, 2021, pp. 438–443, doi: 10.1109/SmartGridComm51999.2021.9631995.
[37] J. Stiasny, G. S. Misyris, and S. Chatzivasileiadis, "Transient stability analysis with physics-informed neural networks," 2021, *arXiv:2106.13638*.
[38] G. Misyris, "Towards zero-inertia power systems: Stability analysis, control & physics-informed neural networks," Ph.D. dissertation, Dep. Elect. Eng., Tech. Univ. of Denmark, 2021.
[39] M. Lahariya, K. Farzaneh, C. Develder, and G. Crevecoeur, "Physics-informed recurrent neural networks for the identification of a generic energy buffer system," in *Proc. 10th IEEE Data-Driven Control Learn. Syst. Conf.*, 2021, pp. 1–6.
[40] G. S. Misyris, J. Stiasny, and S. Chatzivasileiadis, "Capturing power system dynamics by physics-informed neural networks and optimization," in *Proc. 60th IEEE Conf. Dec. Cont. (CDC)*, 2021, pp. 4418–4423, doi: 10.1109/CDC45484.2021.9682779.
[41] Z. Ping, Z. Yin, X. Li, Y. Liu, and T. Yang, "Deep Koopman model predictive control for enhancing transient stability in power grids," *Int. J. Robust Nonlinear Control*, vol. 31, no. 6, pp. 1964–1978, 2021.
[42] P. You, J. Pang, and E. Yeung, "Deep Koopman controller synthesis for cyber-resilient market-based frequency regulation," *IFAC-PapersOnLine*, vol. 51, no. 28, pp. 720–725, 2018.
[43] J. Sun *et al.*, "An integrated critic-actor neural network for reinforcement learning with application of DERs control in grid frequency regulation," *Int. J. Elect. Power Energy Syst.*, vol. 111, pp. 286–299, 2019.
[44] E. Yeung, S. Kundu, and N. Hodas, "Learning deep neural network representations for Koopman operators of nonlinear dynamical systems," in *Proc. IEEE Amer. Control Conf.*, 2019, pp. 4832–4839.
[45] R. Yousefian and S. Kamalasadan, "Energy function inspired value priority based global wide-area control of power grid," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 552–563, Mar. 2018.
[46] X. S. Zhang, T. Yu, Z. N. Pan, B. Yang, and T. Bao, "Lifelong learning for complementary generation control of interconnected power grids with high-penetration renewables and EVs," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4097–4110, Jul. 2018.
[47] L. Yin, T. Yu, L. Zhou, L. Huang, X. Zhang, and B. Zheng, "Artificial emotional reinforcement learning for automatic generation control of large-scale interconnected power grids," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 9, pp. 2305–2313, 2017.
[48] X. Hu, H. Hu, S. Verma, and Z.-L. Zhang, "Physics-guided deep neural networks for power flow analysis," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2082–2092, May 2021.
[49] V. Bolz, J. Ruess, and A. Zell, "Power flow approximation based on graph convolutional networks," in *Proc. 18th IEEE Int. Conf. Mach. Learn. Appl.*, 2019, pp. 1679–1686.
[50] B. Donon, B. Donnot, I. Guyon, and A. Marot, "Graph neural solver for power systems," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2019, pp. 1–8.
[51] B. Donon, R. Clement, B. Donnot, A. Marot, I. Guyon, and M. Schoenauer, "Neural networks for power flow: Graph neural solver," *Electric Power Syst. Res.*, vol. 189, 2020, Art. no. 106547.
[52] F. Fioretto, T. W. Mak, and P. Van Hentenryck, "Predicting ac optimal power flows: Combining deep learning and lagrangian dual methods," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, pp. 630–637.
[53] L. Zhang, Y. Chen, and B. Zhang, "A convex neural network solver for DCOPF with generalization guarantees," *IEEE Trans. Control Netw. Syst.*, to be published, doi: 10.1109/TCNS.2021.3124283.
[54] M. K. Singh, S. Gupta, V. Kekatos, G. Cavraro, and A. Bernstein, "Learning to optimize power distribution grids using sensitivity-informed deep neural networks," in *Proc. IEEE Int. Conf. Commun. Control Comput. Technol. Smart Grids*, 2020, pp. 1–6.
[55] R. Nellikkath and S. Chatzivasileiadis, "Physics-informed neural networks for minimising worst-case violations in dc optimal power flow," in *Proc. IEEE Int. Conf. Commun., Cont., Comput. Technol. Smart Grids (SmartGridComm)*, 2021, pp. 419–424, doi: 10.1109/SmartGridComm51999.2021.9632308.
[56] R. Nellikkath and S. Chatzivasileiadis, "Physics-informed neural networks for ac optimal power flow," 2021, *arXiv:2110.02672*.
[57] S. Gupta, V. Kekatos, and M. Jin, "Deep learning for reactive power control of smart inverters under communication constraints," in *Proc. IEEE Int. Conf. Commun. Control Comput. Technol. Smart Grids*, 2020, pp. 1–6.
[58] Z. Yan and Y. Xu, "Real-time optimal power flow: A Lagrangian based deep reinforcement learning approach," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 3270–3273, Jul. 2020.
[59] X. Sun and J. Qiu, "Two-stage volt/var control in active distribution networks with multi-agent deep reinforcement learning method," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 2903–2912, Jul. 2021.
[60] D. Cao *et al.*, "Deep reinforcement learning enabled physical-model-free two-timescale voltage control method for active distribution systems," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 149–165, Jan. 2022.
[61] D. Cao *et al.*, "Data-driven multi-agent deep reinforcement learning for distribution system decentralized voltage control with high penetration of PVs," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 4137–4150, Sep. 2021.
[62] Y. Zhang, F. Qiu, T. Hong, Z. Wang, and F. F. Li, "Hybrid imitation learning for real-time service restoration in resilient distribution systems," *IEEE Trans. Ind. Inform.*, vol. 18, no. 3, pp. 2089–2099, Mar. 2021.
[63] X. Zhou, S. Wang, R. Diao, D. Bian, J. Duan, and D. Shi, "Rethink AI-based power grid control: Diving into algorithm design," in *Proc. Adv. Neural Inf. Process. Syst.: Workshop: Mach. Learn. Eng. Model., Simul. Design*, Dec. 2020.
[64] W. Li and D. Deka, "Physics-informed learning for high impedance faults detection," in *Proc. IEEE Madrid PowerTech*, 2021, pp. 1–6.
[65] K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He, "Fault location in power distribution systems via deep graph convolutional networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 1, pp. 119–131, Jan. 2019.
[66] W. Li and D. Deka, "Physics-informed graph learning for robust fault location in distribution systems," 2021, *arXiv:2107.02275*.
[67] M. Liang, Y. Meng, J. Wang, D. L. Lubkeman, and N. Lu, "FeederGAN: Synthetic feeder generation via deep graph adversarial nets," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1163–1173, Mar. 2021.
[68] M. Khodayar and J. Wang, "Deep generative graph learning for power grid synthesis," in *Proc. Int. Conf. Smart Energy Syst. Technol.*, 2021, pp. 1–6.
[69] J. Park and J. Park, "Physics-induced graph neural network: An application to wind-farm power estimation," *Energy*, vol. 187, 2019, Art. no. 115883.
[70] T. Zhang, Z. Li, Q. Wu, and X. Zhou, "Decentralized state estimation of combined heat and power systems using the asynchronous alternating direction method of multipliers," *Appl. Energy*, vol. 248, pp. 600–613, 2019.
[71] S. K. Azman, Y. J. Isbeih, M. S. El Moursi, and K. Elbassioni, "A unified online deep learning prediction model for small signal and transient stability," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4585–4598, Nov. 2020.
[72] M. Netto and L. Mili, "A robust data-driven Koopman Kalman filter for power systems dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 7228–7237, Nov. 2018.
[73] E. King *et al.*, "Solving the dynamics-aware economic dispatch problem with the Koopman operator," in *Proc. 12th ACM Int. Conf. Future Energy Syst.*, 2021, pp. 137–147.
[74] M. Korda, Y. Susuki, and I. Mezić, "Power grid transient stabilization using Koopman model predictive control," *IFAC-PapersOnLine*, vol. 51, no. 28, pp. 297–302, 2018.
[75] Y. Susuki and I. Mezić, "Nonlinear Koopman modes and power system stability assessment without models," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 899–907, Mar. 2014.
[76] E. Barocio, B. C. Pal, N. F. Thornhill, and A. R. Messina, "A dynamic mode decomposition framework for global power system oscillation analysis," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 2902–2912, Nov. 2015.
[77] H. R. Chamorro, C. A. Ordonez, J. -H. F. Peng Gonzalez-Longatt, and V. K. Sood, "Coherency estimation in power systems: A Koopman operator approach," in *Computational Intelligence and Optimization Methods for Control Engineering*. Berlin, Germany: Springer, 2019, pp. 201–225.

[78] C. Bakker, A. Bhattacharya, S. Chatterjee, C. J. Perkins, and M. R. Oster, "The Koopman operator: Capabilities and recent advances," in *Proc. IEEE Resilience Week*, 2020, pp. 34–40.

[79] B. Lusch, J. N. Kutz, and S. L. Brunton, "Deep learning for universal linear embeddings of nonlinear dynamics," *Nature Commun.*, vol. 9, no. 1, pp. 1–10, 2018.

[80] S. Pan and K. Duraisamy, "Physics-informed probabilistic learning of linear embeddings of nonlinear dynamics with guaranteed stability," *SIAM J. Appl. Dyn. Syst.*, vol. 19, no. 1, pp. 480–509, 2020.

[81] B. Huang and J. Wang, "Deep-reinforcement-learning-based capacity scheduling for PV-battery storage system," *IEEE Trans. Smart Grid*, vol. 12, no. 3, pp. 2272–2283, May 2021.

[82] M. Glavic, "(Deep) reinforcement learning for electric power system control and related problems: A short review and perspectives," *Annu. Rev. Control*, vol. 48, pp. 22–35, 2019.

[83] F. Li and Y. Du, "From alphaGO to power system AI: What engineers can learn from solving the most complex board game," *IEEE Power Energy Mag.*, vol. 16, no. 2, pp. 76–84, Mar./Apr. 2018.

[84] F. Arnold and R. King, "State-space modeling for control based on physics-informed neural networks," *Eng. Appl. Artif. Intell.*, vol. 101, 2021, Art. no. 104195.

[85] M. A. Roehrl, T. A. Runkler, V. Brandtstetter, M. Tokic, and S. Obermayer, "Modeling system dynamics with physics-informed neural networks based on lagrangian mechanics," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 9195–9200, 2020.

[86] Y. Xu, R. Wang, L. Yang, A. Singh, and A. Dubrawski, "Preference-based reinforcement learning with finite-time guarantees," *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 18784–18794, 2020.

[87] M. Ghavamzadeh *et al.*, "Bayesian reinforcement learning: A survey," *Found. Trends Mach. Learn.*, vol. 8, no. 5/6, pp. 359–483, 2015.

[88] S. Gros and M. Zanon, "Data-driven economic NMPC using reinforcement learning," *IEEE Trans. Autom. Control*, vol. 65, no. 2, pp. 636–648, Feb. 2020.

[89] L. Brunke *et al.*, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annu. Rev. Control*, *Robot., Auton. Syst.*, 2022, in preparation.

[90] L. Yang, D. Zhang, and G. E. Karniadakis, "Physics-informed generative adversarial networks for stochastic differential equations," *SIAM J. Sci. Comput.*, vol. 42, no. 1, pp. A292–A317, 2020.

[91] A. Venzke and S. Chatzivasileiadis, "Verification of neural network behaviour: Formal guarantees for power system applications," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 383–397, Jan. 2021.

[92] Y. Liu, B. Xu, A. Botterud, N. Zhang, and C. Kang, "Bounding regression errors in data-driven power grid steady-state models," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 1023–1033, Mar. 2021.

[93] Y. Liu, N. Zhang, Y. Wang, J. Yang, and C. Kang, "Data-driven power flow linearization: A regression approach," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2569–2580, May 2019.

[94] J. Yu, Y. Weng, and R. Rajagopal, "Robust mapping rule estimation for power flow analysis in distribution grids," in *Proc. North Amer. Power Symp.*, 2017, pp. 1–6.

[95] B. Recht, "A tour of reinforcement learning: The view from continuous control," *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 2, pp. 253–279, 2019.

[96] B. Stott and O. Alsac, "Fast decoupled load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 859–869, May 1974.

[97] S. H. Low, "Convex relaxation of optimal power flow-part I: Formulations and equivalence," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 15–27, Mar. 2014.

[98] M. Sadoughi and C. Hu, "Physics-based convolutional neural network for fault diagnosis of rolling element bearings," *IEEE Sensors J.*, vol. 19, no. 11, pp. 4181–4192, Jan. 2019.

[99] B. Bahmani and W. Sun, "Training multi-objective/multi-task collocation physics-informed neural network with student/teachers transfer learnings," 2021, *arXiv:2107.11496*.

[100] Y. Yu, A. Srivastava, and S. Canales, "Conditional LSTM-GAN for melody generation from lyrics," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 17, no. 1, pp. 1–20, 2021.

[101] A. Lukashevich and Y. Maximov, "Power grid reliability estimation via adaptive importance sampling," *IEEE Control Syst. Lett.*, vol. 6, pp. 1010–1015, 2022.

[102] T. Ma, D. A. Barajas-Solano, R. Tipireddy, and A. M. Tartakovsky, "Physics-informed Gaussian process regression for probabilistic states estimation and forecasting in power grids," 2020, *arXiv:2010.04591*.

[103] M. Jalali, V. Kekatos, S. Bhela, and H. Zhu, "Inferring power system frequency oscillations using gaussian processes," in *Proc. IEEE Conf. Decis. Control*, Austin, TX, USA, 2022.

[104] X. Lei, Z. Yang, J. Yu, J. Zhao, Q. Gao, and H. Yu, "Data-driven optimal power flow: A physics-informed machine learning approach," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 346–354, Jan. 2021.

[105] D. Silver *et al.*, "Mastering the game of go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, 2017.

[106] V. Nair *et al.*, "Solving mixed integer programs using neural networks," 2020, *arXiv:2012.13349*.

[107] H. Zhang, T.-W. Weng, P.-Y. Chen, C.-J. Hsieh, and L. Daniel, "Efficient neural network robustness certification with general activation functions," *Adv. Neural Inf. Process. Syst.*, vol. 31, pp. 4939–4948, 2018.

[108] A. Vaswani *et al.*, "Attention is all you need," in *Proc. Adv. neural Inf. Process. Syst.*, 2017, pp. 5998–6008.

[109] J. Wu, S. Tang, C. Huang, D. Zhang, and Y. Zhao, "Review of attention mechanism in electric power systems," in *Proc. Int. Conf. Artif. Intell. Secur.*, 2021, pp. 618–627.

[110] E. Farhi *et al.*, "Classification with quantum neural networks on near term processors," *Quantum Rev. Lett.*, vol. 1, no. 2, pp. 10–37686, 2020.

[111] S. Lloyd and C. Weedbrook, "Quantum generative adversarial learning," *Phys. Rev. Lett.*, vol. 121, Jul. 2018, Art. no. 040502.

[112] G. Verdon, T. McCourt, E. Luzhnica, V. Singh, S. Leichenauer, and J. Hidary, "Quantum graph neural networks," 2019, *arXiv:1909.12264*.

[113] J. Bongard and H. Lipson, "Automated reverse engineering of nonlinear dynamical systems," *Proc. Nat. Acad. Sci.*, vol. 104, no. 24, pp. 9943–9948, 2007.

[114] M. Schmidt and H. Lipson, "Distilling free-form natural laws from experimental data," *Science,* vol. 324, no. 5923, pp. 81–85, 2009.

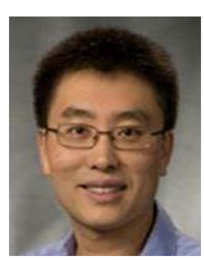
[115] S.-M. Udrescu and M. Tegmark, "Ai feynman: A physics-inspired method for symbolic regression," *Sci. Adv.*, vol. 6, no. 16, 2020, Art. no. eaay2631.

[116] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 248–255.

[117] G. Brockman *et al.*, "Openai gym," 2016, *arXiv:1606.01540*.

[118] S. Babaeinejadsarookolaee *et al.*, "The power grid library for benchmarking ac optimal power flow algorithms," 2019, *arXiv:1908.02788*.

Bin Huang (Student Member, IEEE) received the B.S. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2016, and the M.S. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2019. He is currently working toward the Ph.D. degree in electrical engineering with Southern Methodist University, Dallas, TX, USA. His research interests include the area of machine learning, deep learning, decision making methods, and their applications on smart gird.



Jianhui Wang (Fellow, IEEE) is currently a Professor with the Department of Electrical and Computer Engineering, Southern Methodist University, Dallas, TX, USA. His research interests include smart grid, microgrids, power system operation and control, renewable integration, grid resilience, and cybersecurity.

Dr. Wang is the past Editor-in-Chief of the IEEE TRANSACTIONS ON SMART GRID and an IEEE PES Distinguished Lecturer. He is also a Guest Editor of a Proceedings of the IEEE special issue on power grid resilience. He was the recipient of the IEEE PES Power System Operation Committee Prize Paper Award in 2015, 2018 Premium Award for Best Paper in *IET Cyber-Physical Systems: Theory & Applications*, Best Paper Award in IEEE TRANSACTIONS ON POWER SYSTEMS in 2020, and IEEE PES Power System Operation, Planning and Economics Committee Prize Paper Award in 2021. Dr. Wang is a Clarivate Analytics highly cited Researcher for production of multiple highly cited papers that rank in the top 1% by citations for field and year in Web of Science during 2018–2021.